NEW YORK CITY COLLEGE OF TECHNOLOGY
COMPUTER SYSTEMS TECHNOLOGY DEPARTMENT
COURSE OUTLINE
CST 3604 – DESIGN OF DISTRIBUTED DATABASES

COURSE DESCRIPTION
This course is the continuation of the course “Database Design”. The course concentrates on: the physical design and implementation of databases; functionality of Database Management Systems (DBMS) in support of concurrent, secure, well-performing, reliable, scalable database solutions. Also discussed are: special database architectures and topics – data warehouses and business intelligence, and semi-structured stores (XML and Cloud). Database and data warehouse concepts are illustrated on Oracle RDBMS, XML and Cloud stores are illustrated on various open-source products.

4 Class Hours, 3 Credits

LEARNING OUTCOMES
After finishing the course, the students should be able to:

• Understand the goals of physical data design and implementation of a database. Design the physical model of a database in Oracle DBMS, including a distributed solution.
• Understand the basic features of DBMSs for support of query processing, transaction management, concurrent access, security, reliability
• Participate in resolving the goals of physical design of a database and understand the approaches to implementing user requirements on performance, security, consistency, and reliability with the help of a particular DBMS.
• Implement basic support of database security, performance improvement, reliability, and consistent transactions.
• Understand the purpose and architecture of special databases: data warehouses, semi-structured, Cloud
• Design and implement simple data warehouses and semi-structured stores

ASSESSMENT CRITERIA
Students will be assessed in exams, homework, case assignments, and through class participation. The major areas include:

• The basic understanding of the physical data model and features of a DBMS for its support.
• Implementing, supporting and maintaining a database (in Oracle).
• Designing and implementing a simple data warehouse
• Designing and implementing a simple semi-structured store

General Education Outcomes

• SKILLS/Inquiry/Analysis: Students will employ scientific reasoning and logical thinking.

• SKILLS/Communication: Students will communicate in diverse settings and groups, using written (both reading and writing), oral (both speaking and listening), and visual means

• VALUES, ETHICS, RELATIONSHIPS / Professional/Personal Development: Students will work with teams, including those of diverse composition. Build consensus. Respect and use creativity.

PREREQUISITES
Completion of CST3504 with the grade C or higher.

REQUIRED TEXTBOOK
3. Oracle 11g Documentation, http://www.oracle.com/pls/db111/homepage. The books:
   a. Application Development - Essentials
b. Application Development – Concepts  
d. Database Administrator’s Guide. Part V. Distributed Database Management  
e. Performance.  

**Attendance Policy:**  
You are permitted to be absent from a class a maximum of three class sessions. This is in accordance with the college policy that sets the maximum number of permissible absences at 10% of the number of class meetings scheduled for the semester.  

**Academic Integrity Policy:**  
The instructor of the course has the authority to give a grade of F if the student submits the work of another person in a manner that represents his/her work, or knowingly permits one’s work to be submitted by another person without the instructor’s permission. All class projects must be on your own floppy disk.  
For further information see Student Handbook.  

**Tests:**  
- Final Exam 30%  
- Midterm Exam 30%  
- Tests & Quizzes 15%  
- Case Assignments 15%  
- Homework 10%  

**Case Assignments**  
Case assignments include physical design of the centralized and distributed databases for a particular case (for this purpose the cases from Appendix 1 of the textbook can be used). Design of the distributed database must include primary and derived horizontal fragmentation, vertical fragmentation, allocation and replication of fragments, and semantic control. The distribution of the database must be transparent to users.

On the local databases of the distributed solution and on the global distributed database the students must implement the basic security measures, define queries for the users’ requests and suggest the measures for improving performance of these queries, and design a business transaction ensuring its consistency.

The assignments must be implemented in Oracle DBMS.  

**Grading Policy:**  
You cannot get a passing grade unless all case assignments are completed. The professor preserves the right to ask you to defend any of your case assignments or tests.  

**COURSE OUTLINE**  
*Chapters of [2] are referenced directly; for the chapters from [3], the book of the documentation is referenced by the letter under which it is mentioned in the list of required textbooks, e.g. “2c- Chapter 1” means ”Chapter 1 from the book ‘Design and Tuning for Performance’ from Oracle Documentation.*  

<table>
<thead>
<tr>
<th>Week</th>
<th>Subject</th>
<th>Source</th>
</tr>
</thead>
</table>
| 1.   | The physical data model  
Goals of physical modeling  
Using features of the DBMS for physical data design  
Place of the physical design in the database life cycle  
Implementation of the database  
Review of the process of design of relational databases  
Operations of relational algebra | Chapter 1  
Appendix 2 |
| 2.   | The physical data model  
Indices  
Data storage organization  
Heap and organized data storage | Chapter 2 |
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.</td>
<td>Distributed database design</td>
<td>Distributed storage of data. Benefits and problems of design and implementation of distributed databases. Transparency of the physical model. 3-tier database architecture. The physical data model in Oracle. Oracle distributed database.</td>
</tr>
<tr>
<td>5.</td>
<td>Distributed database design</td>
<td>Semantic control.</td>
</tr>
<tr>
<td>13.</td>
<td>Transaction management in distributed database</td>
<td>Transaction management in Oracle. Locking, multiversionsing, different isolation levels, and additional tools. Practicing building transactions and analyzing results of concurrent execution of different simple transactions.</td>
</tr>
</tbody>
</table>
15. Review
Final test